

N63

17207

UNCLASS

CODE 1

2

1

## A SYSTEM APPROACH TO RELIABILITY TESTING

By

Catherine Dryden Mock  
Office of Human Space Flight  
National Aeronautics and Space Administration  
Washington, D. C.

### INTRODUCTION

In recent years a number of papers have been written addressed to the problem of determining system confidence limits for the true system reliability based on available component test results. At the same time, however, little has been developed in the literature concerning the related problem of determining the sample sizes needed for a system's component tests which will demonstrate an adequate level of system reliability. This latter problem is faced by a system contractor who must develop a reliability test program for a complex system, and is the one which will be discussed here.

Generally, the procedure has been to set a system goal and then allocate this value among the various subsystems, components and parts in some rational manner. The numbers thus obtained then become the requirements for the lower level elements. Several questions can be raised by this procedure. First, examining the number selected as the system goal, one might ask whether it represents the desired objective or goal, an expected value, or a lower reliability limit which one desires to exceed with a certain fairly high probability. Since it is frequently not clear which of these interpretations is to be placed on the system goal, it is equally unclear which interpretation applies to the values assigned to the system elements. The problem of demonstrating reliability through testing depends, of course, on the interpretation selected.

Presented at the 1962 Reliability and Statistical Methods in Industry Course, University of California, Los Angeles, August 7, 1962.

100-97617-168 Rev. 1

- 2 -

In the discussion following, the system "goal" is expressed in terms of a lower reliability value, or confidence limit, which one desires to exceed with high probability. This definition applies to the sub-tier element values as well. The problem now reduces itself into determining what these sub-tier element reliability confidence limits should be, so that they are related in a probabilistic sense to the system confidence limit for reliability at a given system confidence level. Once these limits are defined, the component sample sizes for demonstrating system reliability can then be developed.

Before considering a series of examples which examine various alternatives in minimizing system reliability testing and the ideas just expressed, a few comments are in order on the kinds of tests which can be considered as reliability tests. In a general sense, all testing represents some aspect of reliability testing, in that it provides knowledge about the functioning of the system or its elements, which can, in turn, be interpreted in the form of a reliability statement. For purposes of ~~demonstrating~~ reliability, however, one is concerned with tests which provide knowledge of the system's performance during the mission lifetime and under operating conditions and environments; hence tests performed under environments below those required for the mission are of doubtful use for reliability demonstration (unless they appear useful). This would be so if a relationship could be developed which would show the effect of increased stress on reliability. In ~~some~~ ~~most~~ cases the exponential function provides the relationship of operating time to reliability. Unfortunately, it is likely that more tests would be required to develop such a relationship than would be required to demonstrate the reliability at the operating level.

4

3

It is noted, however, that should no failures occur in excess stress tests, the tests can be combined with reliability demonstration tests at the operating level, with the knowledge that the true reliability will be higher than the value demonstrated by numbers of tests. How much higher, of course, would again be predicated on the relationship discussed above.

Finally, since the examples will explore the concept of minimum testing consistent with a specified amount of assurance, some mention should be made of the cost of reliability testing. A thorough discussion of costs has been purposely omitted, since this would add another dimension to an already complex problem. Certainly, concurrently with an analysis such as that to be developed, test costs could also be developed; trade-offs between test costs for various levels of assurance and the cost of achieving the program objectives can be studied before finally arriving at a suitable reliability demonstration test program.

#### Test Plans for the System as a Whole

To illustrate the sequence of steps in developing a reliability demonstration test program, it will be assumed that need exists for an earth orbiting satellite which is required to have a functional reliability of .90 with 90% confidence, for a minimum of 3 months in orbit, for the particular experiment to be conducted. It is further assumed that this satellite, although requiring some newly designed equipment, is still within the existing state of the art with respect to its parts and components. It is planned to demonstrate the above requirements by means of ground tests in an environment simulating operating conditions.

Since it is required that the satellite, itself, perform trouble free over a minimum time (i.e., no failures) with a given assurance, the first

alternative to consider is developing a test program for the satellite system as a whole.

Two types of reliability test plans which can be utilized for this purpose are discussed below: test plans based on the exponential function to failure, and test plans based on the binomial function in which tests can be made for the specified minimum time.

To use test plans based on the exponential distribution, one must assume that times to failure for this type of equipment are random variables with an exponential probability density, verification of the validity of this assumption being made on the basis of previous experience. The exponential assumption implies that the conditional probability in a given time interval is approximately proportional to the length of the time interval, the condition being that the system is operating at the beginning of the interval. This mean time to failure (MTTF) of the satellite associated with the above requirements is then obtained through the reliability function:

$$R(t) = e^{-\frac{t}{MTTF}} \quad (1)$$

If  $R(t)$  = probability of success, (no failure before time  $t$ ),

$t$  = required minimum time in orbit, namely, 3 months or 2,160 hours, and

$\theta$  = mean time to failure (MTTF),

one can substitute these values into (1) to obtain:

$$\theta = 21,600 \text{ hours} \quad (2)$$

To demonstrate this MTTF for the satellite system with the stated assurance of 90%, test plans of the QAB Quality Control and Reliability Handbook N-100, Sampling Procedures and Tables for Life and Reliability

- 4 -

5

6

- 3 -

Testing (Based on Exponential Distribution)<sup>1</sup> can be used. The requirement of 90% confidence that the minimum mean life of 21,600 hours be met is denoted by  $\beta = .10$  and  $\theta_1 = 21,600$  in the notation of this Handbook.

However, since one must enter the tables with  $\theta_0$  or the mean life which will be accepted with high probability (1- $\alpha$ ), we must first compute  $\theta_0$  from the Handbook table 2A-1 for an  $\alpha$  value chosen from those given (i.e.,  $\alpha = .01, .05, .10, .25$  and .50). Examination of the Handbook for plans which provide the simplest and least amount of testing for the requirements given discloses that tests terminated at a pre-assigned time and providing for acceptance if no failure occurs before the termination time meet this requirement (i.e., H-108, Section 2C Plans A-1 through E-1).

Since these plans allow for rejection as soon as a failure occurs, the advantage of sequential testing is gained without having developed the more intricate structure of such tests.<sup>2</sup>

Computations of the required testing for the example cited above follow for Sampling Plan Code C-1 ( $\alpha = \beta = .10$ , i.e., both Producer's and consumer's risks are 10% for  $\theta_0$  and  $\theta_1$ ). Given  $\theta_1 = 21,600$ , and using the value of  $\theta_1/\theta_0 = .046$ , from H-108, Table 2A-1, one finds for this sampling plan:

$$\frac{21,600}{\theta_0} = .046, \quad (3)$$

or  $\theta_0 = 469,565$  hours.

From H-108 Table 2C-1(c) for sample size 2 and rejection number,  $r_1$ , of 1, and using  $\theta_0$  of 469,565, the termination time,  $T$ , is found as follows:

$$\frac{T}{\theta_0} = .053, \quad (4)$$

or  $T = .053 \times 469,565 = 24,887$  hours.

- 5 -

Similarly, for sample size 20, rejection number 1, and  $\theta_0 = 469,565$  hours, H-108 Table 2C-1(c) provides

$$\frac{T}{\theta_0} = .005, \quad (5)$$

or  $T = .005 \times 469,565 = 2,348$  hours.

It is of interest to note that this method will provide approximately the same amount of test time (i.e., "terminal" time,  $T$ ) for the sample sizes selected above no matter which of plans A-1, B-1, C-1, D-1 or E-1 is chosen. The exact solution of the minimum total test time required without failure is given by Equation 3.

To illustrate: Using exponential test plans and the requirement to demonstrate a minimum life of 3 months for a satellite having .90 reliability with 90% confidence with no failures, one must require a minimum mean life of 30 months ( $\theta_1 = 11,420$  hours). The sampling plan which assures this requires development of a satellite which will have a  $\theta_0$  of approximately 600 months or 50 years in order to insure the small risk, 10%, of being rejected. The test time,  $T$ , is approximately 35 months for 2 systems, and 31 months for 20 systems.

Examining the requirement from the point of view of attributes testing for a fixed time, and using test plans similar to MIL-STD-105B<sup>4</sup>, one requires tests of 22 systems over a 3-month period to provide 90% assurance that the material represented is .90 effective. For a given confidence, it can be shown that attribute test plans for sample size,  $n$ , and a fixed minimum test time,  $t^*$ , are equivalent to exponential test plans where the termination time,  $T$ , for a single exponential test unit is

$$\frac{T}{\theta_0} = n t^*, \quad T = n t^*. \quad (6)$$

- 6 -

7

8

The attributes test plan has the advantage of not requiring any assumption about the distribution.

The choice between the two kinds of plans, given that the exponential assumption is valid, is merely dependent on whether time or satellites are more costly. If one must limit funds to the purchase of only 2 satellites for test one must be prepared to wait nearly 3 years to complete the test. If one can afford to purchase 22 satellites for test, one can complete the test in 3 months. Of course, there are intermediate steps along the way, i.e., 14 years of testing 4 satellites, and so on.

With either type of test plan, if a failure occurs during tests the satellite will be rejected. Obviously, the system contractor must then search for the cause of the failure, correct it on all the systems, and repeat that portion of the test in which the previous system failed.

The above testing approach indicates that for long life space systems one cannot afford to demonstrate reliability to a high degree of confidence by applying the well-known sampling test plans to complete systems, without expending large amounts of time or dollars, or both, on the program. In order to overcome the problems encountered with testing complete systems one must necessarily turn to component or subsystem testing.

#### Test Plans for Components

Since it was stated earlier that sampling plans for systems requiring no failures for a certain minimum time can be based on either the exponential distribution (if the assumption holds) or the binomial distribution by satisfying (6), the remainder of the paper will deal with binomial probabilities. For the binomial test plans of components to have useful application to

system performance, all component tests will be made under operational and environmental conditions approximating those met in system use for the minimum time the component is required to function to assure system functioning for its minimum time. This condition applies to the following discussion of component testing without specific mention.

Consider, first, a system composed of two dissimilar components or black boxes in series as represented below.



Let  $q$  be the failure rate of Component A and  $r$  be the failure rate of Component B. By Mood<sup>5</sup> and Bushler<sup>6</sup> it can be shown that if

$$P[(1-q)] \geq (1-c_1)] \geq 1-\alpha, \quad (7)$$

and

$$P[(1-r)] \geq (1-c_2)] \geq 1-\alpha, \quad (8)$$

where  $(1-c_1)$  and  $(1-c_2)$  represent the lower confidence limits for the success rates  $(1-q)$  and  $(1-r)$  of Components A and B respectively, with  $(1-\alpha)$  confidence coefficient (where  $\alpha$  is a small probability), then

$$P[(1-q)(1-r) \geq (1-c_1)(1-c_2)] \geq (1-\alpha)^2. \quad (9)$$

Assuming equal sample sizes,  $n$ , and no failures for both components, the lower confidence limits for systems made up of Components A and B can be expressed as

$$(1-c_1)(1-c_2) = \alpha^{\frac{2}{n}} \quad (9)$$

with  $(1-\alpha)^2$  confidence coefficient.

For a system of two components in series circuit to have a lower confidence limit on reliability of .90 with 90% confidence and the above

assumption, one can solve (9) for  $n$  as follows:

$$.90 = (.05)^{\frac{2}{n}}$$

(10)

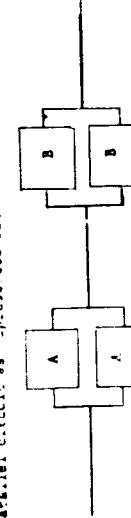
or

$$n = 54.$$

(since  $(.49)^2 = .90$  implies that  $\alpha \approx .05$ )

Comparing this amount of testing with that required for the system as a whole, computed earlier at 22, it is apparent that this method more than doubles the amount of testing required to achieve the necessary system confidence. Buchler states, however, that the solution obtained in this manner is "not a very good one as the confidence intervals are not as small as possible. It is interesting to note that for the particular series case illustrated, i.e., equal sample sizes and no failures allowed, it can be shown, using another approach discussed by Buchler, that the sample sizes required for component testing are exactly the same as that required for the system itself. This is true no matter how many components there are in series.

Consider, next, a system composed of two dissimilar components in a series-parallel circuit as represented below.



Again, let  $q$  be the failure rate of  $A$  and  $r$  be the failure rate of  $B$ . By Stuck<sup>7</sup>, it can be shown that the probability of failure of the system is:

$$P(q,r) = 1 - (1-q^2)(1-r^2).$$

Maximizing this function, subject to the restriction  $(1-q)^n(1-r)^m = 1 - \beta$ , where  $(1 - \beta)$  is the confidence coefficient ( $\beta$  is a large probability), results in the upper confidence limit for system failures,  $C(n,n,0,0,\beta)$ ,

- 9 -

assuming equal sample sizes,  $n$ , and no failures for both components. This upper confidence limit can be expressed as:

$$C(n,n,0,0,\beta) = L1 - (1-\beta)^{\frac{1}{n}}.2. \quad (12)$$

For a system of two components in series - parallel circuit to have an upper confidence limit on system failure of .10 (or for success a lower limit of .90) with 90% confidence and the above assumptions, one can solve (12) for  $n$  as follows:

$$.10 = L1 - (1-.90)^{\frac{1}{n}}.2 \quad (13)$$

or

$$n = 6.$$

The value of  $n$  for this case can also be read directly from the lowest curve on the charts provided on Figures (4), (5) and (6). An

upper confidence limit to .5.

reference<sup>7</sup> for confidence levels of 90%, 95% and 99% and any choice of

upper confidence limits to .5.

It is further shown by Stuck that the relationship (12) holds for the  $k$ -fold series-parallel circuit, where  $k$  is the number of dissimilar components in this configuration, provided that equal sample sizes are taken for each type of component and no failures are recorded. Hence,

in comparison with usual series circuits, the value of parallelizing series circuits from the point of view of saving in component testing is obvious. This important conclusion by Stuck does not yet appear to have been sufficiently generally appreciated.

For more complex circuits, Stuck merely outlines the approach of constructing the upper confidence limits for system failure by means of computing the probabilities associated with all possible orderings of the sample points (failures) for each component. Even with a computer,

- 10 -

11

12

this may become a rather imposing task if the number of components is large. In this connection, Madansky<sup>8</sup> gives an approximate solution for any series, parallel, or series-parallel system which is relatively easy to perform; however, for cases of extremely high reliability, the author states the approximation is poor.

Two other references are of interest. Karniol and Yousteff<sup>9</sup>

provide a method for obtaining confidence limits on the system based on the use of the upper confidence limits for failure rates. This represents an extreme case approach, and will tend to give a lower bound on system reliability which is inefficient when compared with specific solutions such as discussed above. Rosenblatt and Glinzko<sup>10</sup> discuss confidence limits for a number of simple functions for which known distribution theory is available.

To overcome some of the limitations noted above, an alternate general approach for determining test plans for components based on a given probability level for the system's reliability is introduced below.

#### A General Approach for Determining Component

##### Test Plans Based on the System Model

The mathematical model of the system, developed to assess the inherent reliability of the system<sup>11</sup>, can be transformed into an expression which provides, with adequate realism for this purpose, a value for the probability of a system performing successfully for the minimum specified time in the following general way:

$$P_S = f(P_1, P_2, \dots, P_N, \dots, P_M). \quad (14)$$

In this equation,  $P_i$  is the probability of system success and  $P_i$  is the probability of successful functioning of the  $i$ th item.

- 11 -

It is assumed that the logical diagram for over-all system operation leading to equation (14) can always be constructed by proper choice of functional blocks for the individual items, each of these consisting of appropriate groupings of subsystems, components, and/or parts, and such that the  $P_i$  are functionally independent variables. It should be chosen so that, for a given system and computer, the analysis to be described below constitutes a reasonable programming and running problem. It is clear that each item can be functionally subdivided for further analyses in the same way that equation (14) performs this subdivision for the system as a whole.

It is assumed that one can obtain values,  $P_i$ , for the state-of-the-art capabilities for each item to perform its required functions satisfactorily while the system is operative for the minimum specified time. Let  $P$  be the value  $P_0$  is to exceed with the probability  $(1-\alpha)$ , where  $\alpha$  is some small probability, and let the  $n_i$  be the "zero defective" sample sizes which will assure this condition. This provides the minimum sample size for each component, and is not a necessary restriction. The problem is to determine  $P$  and  $n_i$  knowing the function  $f$  given by (14).

A set of lower limits is defined which one is generally willing to consider for the  $P_i$  by the symbols  $P_i^-(n)$ . One can then divide each of the intervals  $1-P_i^-(n)$  into some convenient number of intervals, say 100, generating this number of distinctive values of  $P_i$ ,  $\gamma_i$  satisfying the inequalities:

$$\begin{aligned} P_i^-(n) &\leq P_i, \gamma_i < 1 & i = 1, \dots, N \\ \gamma_i &= 1, \dots, 100. \end{aligned} \quad (15)$$

If one doesn't know the distributions of the  $P_i$ ,  $\gamma_i$  one can assume that all of them are equally likely. If one knows the distributions, one has

- 12 -

13

14

a measure of the relative frequency of occurrence of each  $P_i \gamma_i$  for a given value of  $i$ . In either case, one can use a table of random numbers to select a  $\gamma_i$  for each  $i$  thereby obtaining a first set of the  $P_i \gamma_i$ , as follows:

$$P_1 \cdot \frac{1}{\gamma_1}, P_2 \cdot \frac{1}{\gamma_1}, \dots, P_1 \cdot \frac{1}{\gamma_1}, \dots, P_N \cdot \frac{1}{\gamma_N} \quad (14)$$

where  $\gamma_1$  defines the 1st value of  $P_1 \gamma_1$  selected from the set (15).

Inserting the value from (16) into (16) one can then compute an estimate  $P_0^1$  of  $P_0$ :

Repeating the process, a very large number of times, say  $Q$ , one will have obtained  $Q$  estimates of  $P_0$  as follows:

$$\begin{aligned} P_0^1 &= f(P_1 \cdot \frac{1}{\gamma_1}; P_2 \cdot \frac{1}{\gamma_1^2}; \dots; P_1 \cdot \frac{1}{\gamma_1^2}; \dots; P_N \cdot \frac{1}{\gamma_N^2}) \\ P_0^2 &= f(P_1 \cdot \frac{2}{\gamma_1}; P_2 \cdot \frac{2}{\gamma_1^2}; \dots; P_1 \cdot \frac{2}{\gamma_1^2}; \dots; P_N \cdot \frac{2}{\gamma_N^2}) \\ &\vdots \\ P_0^Q &= f(P_1 \cdot \frac{Q}{\gamma_1}; P_2 \cdot \frac{Q}{\gamma_1^2}; \dots; P_1 \cdot \frac{Q}{\gamma_1^2}; \dots; P_N \cdot \frac{Q}{\gamma_N^2}) \end{aligned} \quad (17)$$

$$P_0^Q = f(P_1 \cdot \gamma_1^Q; P_2 \cdot \gamma_2^Q; \dots; P_1 \cdot \gamma_1^Q; \dots; P_N \cdot \gamma_N^Q).$$

Let  $P_0^M$  and  $P_0^m$  be the smallest and largest values of  $P_0^j$  given by (17).

Then one can divide the interval  $(P_0^m, P_0^M)$  into, say 100, equal parts and construct the corresponding frequency distribution for the  $P_0^j$ , as shown below:

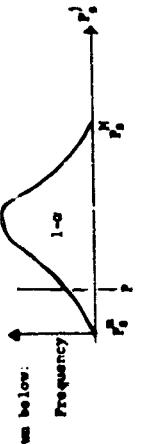


FIG. 1

Using the distribution one can now define a value of  $P$  such that

$$\frac{\text{Number of cases for which } P_0^j \geq P}{Q} = 1-\alpha. \quad (18)$$

# 15

limiting consideration to values of  $P_0^j$  satisfying (18), one can use a similar method to establish the required sample sizes for each item.

Consider, for example, the  $K$  ( $K \leq Q$ ) values of  $P_2, \gamma_2^k$  which are found in the  $K$  values of the  $P_0^j$  satisfying (16). Dividing the interval between the maximum and minimum of these values and constructing a frequency distribution as before, one can find a value  $P_2^*$  such that

$$\frac{\text{Number of cases for which } P_2, \gamma_2^k \geq P_2^*}{K} = 1-\alpha. \quad (19)$$

The resulting value of  $P_2$  can then be employed as a desired lower confidence limit, for a given confidence coefficient  $(1-\alpha)$ , in order to obtain the sample size required if one observes zero defectives, using the binomial theory as given in reference 5. More one must solve the following equation for  $n_2$ :

$$\sum_{n_2=k}^{n_2^*} \binom{n_2}{k} (P_2^*)^{n_2} (1-P_2^*)^{n_2-k} = \alpha. \quad (20)$$

where  $k$  is the number of non-defectives in the sample of  $n_2$ . For the case at hand,  $k = n_2$ , and

$$(P_2^*)^{n_2} = \alpha. \quad (21)$$

or

$$n_2 = \frac{\log \alpha}{\log (P_2^*)}.$$

The above procedure might be considered an application of the so-called "Monte Carlo" technique which has found recent applications in many fields.

## SUMMARY

As indicated in the introduction, the foregoing general method is not the solution to the classical confidence limit problem in which one wishes

- 14 -

(13)

# 16

to determine the system confidence limit for the true system reliability based upon component test results. It develops, in fact, a tolerance limit, since it provides a particular value such that the probability is high that actual system reliabilities will exceed this value. The tolerance limits developed in turn for the sub-tier items are used to determine component sample sizes, which will provide the desired system reliability.

It should be noted that the estimates of a priori distributions for the  $P_i$  may be subject to some error which may, in turn, affect the outcome of the experiment. The consequences of using a priori uniform distributions have not been investigated; however, it is believed that the method provides a conservative estimate of the tolerance limit.

Before one concludes from the above discussion that there is no need for system testing at all, one should recall that there may be system reliability problems which one has initially overlooked in the model. Therefore, it is necessary to build at least one system for reliability demonstration testing, as a check on this accuracy of the model. The system itself should be tested the minimum time it is to function, as a phase of the process of verifying that the system does, in fact, meet the minimum reliability requirement, as demonstrated by component testing.

The implications of the general solution to the reliability testing problem outlined above can be stated as follows:

1. The advantage of keeping the number of various components to a minimum, i.e., making use of the same item in as many system applications as possible, is obvious.
2. The introduction of new components or configurations may require a re-evaluation of the test program underway. The best way to study this is to

reprogram the above experiment using new configurations or component inputs to determine the effect on component testing.

The use of paralleling in series circuits for the system can reduce materially the number of component tests required, resulting in a significant cost saving, as well as increased reliability.

There is no need to test large numbers of complete systems for long periods of time. The need is only for verification of the procedure proposed by testing at least one system. This again represents a substantial saving in test costs.

Finally, from a planning point of view, it is clear that the system contractor must approach the whole area of system design and test for reliability from the viewpoint of system modeling. This means that he must obtain personnel with the necessary skills and qualifications to study reliability analyses and test planning by means of mathematical and physical simulation techniques. Large scale computers are required to perform analyses for highly complex systems. All of this must be accomplished early in the design and development stages to be of the greatest usefulness in the program. Such a probability model approach appears to be the best way to achieve an optimum reliability test program for long life complex space systems.

Except for the introduction, this paper, in substantially the same form, was presented at the 8th National Symposium on Reliability and Quality Control, Washington, D. C. on January 10, 1942.

#### References

1. NASA Quality Control and Reliability Handbook (Incorporated), H-100, *Sample Procedures and Tables for Life and Reliability Testings*.

- 15 -

- 15 -

17

18

- Based on Exponential Distribution, U.S. Government Printing Office, Washington, D.C., 29 April 1940.
2. Sobel, M. and Tsichendorf, J. A., "Acceptance Sampling with New Life Test Objectives," Proceedings of the Fifth National Seminar on Reliability and Quality Control, Institute of Radio Engineers, New York, N.Y., 1959, pp. 108-118.
3. Epstein, Benjamin, "Estimation from Life Test Data," Tachometers, Vol. 2, No. 4, November, 1940, p. 451.
4. MIL-STD-105B, Sampling Procedures and Tables for Tolerance by Attribute, U.S. Government Printing Office, Washington, D.C., 31 December 1950.
5. Mood, A. M., Introduction to the Theory of Statistics, McGraw-Hill, New York, 1950, pp. 223, 234.
6. Suehler, Robert J., "Confidence Intervals for the Product of Two Binomial Parameters," Journal of the American Statistical Association, Vol. 52, December 1957, pp. 482-492.
7. Stech, G. P., "Upper Confidence Limits for the Failure Probability of Complex Networks," Sandia Corporation, Report SC-4133(TM), December 1957.
8. Nedraley, Albert, "Approximate Confidence Limits for the Reliability of Series and Parallel Systems," The Rand Corporation, RM-2552, April 1940.
9. Karelol, Edwin D. and Youstiff, John S., "Technique in the Demonstration of Reliability Attainment," General Electric Missile and Space Vehicle Department, Philadelphia, Pennsylvania, August 26, 1959.
10. Rosenblatt, Joan R., and Grubbs, H., "Confidence Limits for Reliability," NBS Statistical Eng. Lab., W-61-2, May 1961.

- 18 -

19

20

- 17 -